What is claimed is:

1. A network switch for switching transaction requests among a plurality of servers, the network switch being positioned between the plurality of servers and at least one client, comprising:

a parser operable to parse transaction requests to locate one or more selected fields;

a router operable to forward at least portions of the transaction requests to respective servers in the plurality of servers and transaction responses of the respective servers to the transaction requests to respective clients; and

a tag generator operable to generate a tag associated with a selected server in the plurality of servers, the server being operable to provide information requested by a transaction request, whereby, when a subsequent transaction request is received from the client corresponding to the tagged transaction request, the subsequent transaction request includes the tag and, based on the tag, the router forwards the subsequent transaction request to the selected server.

2. The switch of Claim 1, wherein the tag generator is further operable to append the tag to a server tag generated by the selected server.

3. The switch of Claim 1, wherein each of the plurality of servers has a unique server identifier and the tag associated with each server is based on the corresponding unique server identifier.

4. The switch of Claim 1, wherein the tag generator is operable in a tagging mode and is not operable in a digesting mode, and wherein the switch further comprises:

a cache operable to store a plurality of objects corresponding to transaction requests associated with at least one of the plurality of servers, the objects comprising field information in at least one of the selected fields located by and received from the parser;

a digest generator operable to generate a digest based on the field information in at least one selected field of a transaction request, the digest corresponding to a location in the cache where at least one object corresponding to the transaction request is to be stored; and

a cache processor operable to access the plurality of objects in response to communications received from the router.

5. The switch of Claim 4, wherein the digest generator is operable in the digesting mode and is not operable in the tagging mode.

6. The switch of Claim 1, further comprising a decryption processor that decrypts cipher text transaction requests and provides plain text transaction requests to the parser.

7. The switch of Claim 4, further comprising at least one traffic manager located between the network switch and the at least one client and wherein the digest is generated by a hashing function.

8. The switch of Claim 1, wherein the selected fields include at least a universal resource locator and a cookie.

9. The switch of Claim 1, wherein the router includes a current connection table listing active connections between servers and clients.

10. The switch of Claim 4, wherein the plurality of objects in the cache include a plurality of content addresses for specific content and a corresponding hit counter showing a number of instances in a predetermined period of time in which specific content is requested by transaction requests.

11. A method for switching transaction requests, comprising:

receiving, from a first source, a transaction response associated with first source, the transaction response corresponding to at least a first transaction request;

parsing the transaction response to locate at least a first field;

determining a first tag identifying the first source;

appending the first tag to the first field in the transaction response;

reassembling the transaction response; and

forwarding the transaction response to a destination identified by the transaction response.

12. The method of Claim 11, wherein the first field is associated with a server-generated tag, wherein the first tag is an address, and wherein the first tag is derived from field information in the at least a first field..

13. The method of Claim 11, wherein the first source is a first server in a plurality of servers and the destination is a client and further comprising:

receiving the transaction response after the forwarding step;

storing the first tag in the client's memory; and

forwarding a second transaction request to an address associated with the first server, the second transaction request including the first tag.

14. The method of Claim 13, wherein each server in the plurality of servers has a unique identifier and the first tag is based on the unique identifier associated with the first server.

15. The method of Claim 13, further comprising:

receiving the second transaction request from the client;

parsing for the first field in the second transaction request; and

forwarding the second transaction request to the first server based on the first tag.

16. The method of Claim 11, further comprising:

receiving a second transaction request;

parsing the second transaction request for at least the first field;

determining a digest value based on field information in the at least the first field; and

storing selected information corresponding to the second transaction request at an address based on the digest value.

17. The method of Claim 16, wherein the second transaction request is in hypertext transfer protocol, the digest value is generated by a hashing function, and the field information used to determine the digest value is at least one of a universal resource locator and a cookie.

18. The method of Claim 17, wherein the second transaction request is in cipher text and further comprising after the step of receiving the second transaction request and before the step of parsing the second transaction request:

decrypting the second transaction request.

19. The method of Claim 17, wherein storing step comprises:

at least one of incrementing and decrementing a hit counter;

determining if the hit counter at least one of equals or exceeds a predetermined threshold if the hit counter is incremented or at least one of equals or is less than the predetermined threshold if the hit counter is decremented; and

updating a timestamp associated with the stored information.

20. The method of Claim 19, wherein, when the hit counter at least one of equals or exceeds the predetermined threshold, the method further comprises determining a plurality of network addresses associated with content referenced in the second transaction request.

21. The method of Claim 19, wherein, when the hit counter at least one of equals or exceeds the predetermined threshold, the method further comprises directing the second transaction request to a cache server in a plurality of servers.

22. The method of Claim 19, further comprising:

determining whether the second transaction request is a part of an existing connection between an origin server corresponding to content referenced in the second transaction request and a client;

when the second transaction request is part of an existing connection, forwarding the second transaction request to the origin server; and

when the second transaction request is not part of an existing connection and the hit counter at least one of equals or exceeds the predetermined threshold, forwarding the second transaction request to a cache server different from the origin server.

23. The method of Claim 22, the second transaction request is not part of an existing connection and the hit counter exceeds the predetermined threshold and further comprising:

determining whether the second transaction request can be served by a cache server; and

if the second transaction request cannot be served by the cache server, forwarding the transaction request to the origin server.

24. The method of Claim 22, further comprising:

when the hit counter at least one of equals or exceeds the predetermined threshold, transferring content associated with the second transaction request from the origin server to the cache server.

25. A system for switching transaction requests among a plurality of servers, comprising:

an input port for receiving, from a first server in the plurality of servers, a transaction response of the first server, the transaction response corresponding to at least a first transaction request;

means for parsing the transaction response to locate at least a first field;

means for determining a first tag identifying the first server;

means for appending the first tag to the first field in the transaction response;

means for reassembling the transaction response; and

means for forwarding the transaction response to a client identified by the transaction response.

26. The system of Claim 25, wherein the first field is associated with a server-generated tag.

27. The system of Claim 25, further comprising:

a second input port for receiving the transaction response from the forwarding means;

means for storing the first tag in the client's memory; and

means for forwarding a second transaction request to an address associated with the first server, the second transaction request including the first tag.

28. The system of Claim 25, wherein each server in the plurality of servers has a unique identifier and the first tag is based on the unique identifier associated with the first server.

29. The system of Claim 25, wherein the input port receives a second transaction request and further comprising:

means for parsing the second transaction request for at least the first field;

means for determining a digest value based on field information in the at least the first field; and

means for storing selected information corresponding to the second transaction request at an address based on the digest value.

30. The system of Claim 29, wherein the second transaction request is in hypertext transfer protocol, the digest value is generated by a hashing function, and the field information used to determine the digest value is at least one of a universal resource locator and a cookie.

31. The system of Claim 29, wherein the second transaction request is in cipher text and further comprising between the input port and the parsing means:

means for decrypting the second transaction request.

32. A system, comprising:

a communications network;

a plurality of replicated servers connected to the network, all of the replicated servers having a same network address and all of the replicated servers serving the same replicated information, each of the replicated servers being configured to receive a first transaction request associated with an individual transaction and to provide a response to the first transaction request, the response including a first tag that corresponds to the transaction, the first tag being generated by a first replicated server; and

a network switch connecting the replicated servers to the network, the network switch being configured to generate a second tag associated with the first replicated server, to append the second tag to the first tag in the response, and to direct to the first replicated server subsequently received transaction requests including the first and second tags.

33. The system of Claim 32, wherein the network switch is operable to store the first tag and to parse the first transaction request.

34. The system of Claim 33, wherein the network switch is operable to decrypt the first transaction request before the network switch parses the first transaction request.

35. The system of Claim 34, wherein the first tag is part of a plurality of stored objects and the plurality of stored objects correspond to the first transaction request and wherein the plurality of stored objects include a hit counter indicating a frequency of transaction requests for content associated with the first transaction request.

36. A method for providing information from a server to a client, comprising:

receiving a first transaction request requesting first information, the first information referencing at least second and third information;

retrieving the first information;

providing the first information to the client;

determining which of the second and third information has been more frequently requested by clients during a first selected time interval;

retrieving the more frequently requested of the second and third information and/or an address associated therewith;

thereafter receiving a second transaction request from the client requesting the more frequently requested of the second and third information; and

providing the more requested of the second and third information to the client.

37. The method of Claim 36, wherein the first information corresponds to a first hot reference counter, the second information corresponds to a second hot reference counter, and the third information corresponds to a third hot reference counter and wherein in the determining step the second hot reference counter is compared with the third hot reference counter.

38. The method of Claim 37, wherein each of the first, second, and third hot reference counters indicates a number of requests received for the corresponding first, second and third information during the first selected time interval.

39. The method of Claim 38, wherein in the first selected time interval the second hot reference counter is greater than the third hot reference counter and in a second selected time interval different from the first selected time interval the third hot reference counter is greater than the second hot reference counter.

40. The method of Claim 39, wherein in the first selected time interval the second information is requested more frequently than the third information and in the second selected time interval the third information is requested more frequently than the second information.

41. The method of Claim 39, further comprising:

decrypting the first transaction request before the first information is retrieved.

42. A method for configuring stored information in at least one cache server, comprising:

comparing first and second hot reference counters corresponding to first and second information to determine which of the first and second information is more frequently requested; and

storing the more frequently requested of the first and second information in a first location and the less frequently requested of the first and second information in a second location and wherein the first location is more accessible than the second location.

43. The method of Claim 42, wherein the first location is located by the server in a search before the second location.

44. The method of Claim 42, wherein the first information is stored at the first location and further comprising:

retrieving third information associated with the first information, the third information being less frequently requested than the second information;

storing the third information at a third location in close proximity to the first location, the third location being more accessible than the second location.